DIFFERENCES BETWEEN WORD INTELLIGIBILITY AND SENTENCE IDENTIFICATION RESPONSES TO "SYNTHETIC" SENTENCES

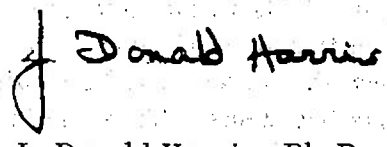
by

Joseph J. Smaldino
with
C. K. Myers

NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY
NAVAL SUBMARINE MEDICAL CENTER REPORT NO. 683

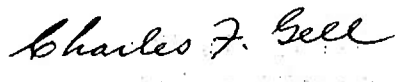
Bureau of Medicine and Surgery, Navy Department
Research Work Unit M4305.08-3003D.04

Transmitted by:



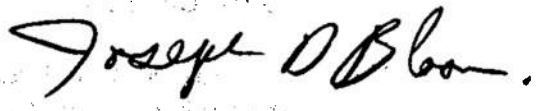
J. Donald Harris, Ph.D.
Head, Auditory Research Branch

Reviewed and Approved by:



Charles F. Gell, M.D., D.Sc.(Med)
Scientific Director
NavSubMedRschLab

Reviewed and Approved by:



J. D. Bloom, CDR MC USN
Officer-in-Charge
NavSubMedRschLab

Approved and Released by:

J. E. Stark, CAPT MC USN
COMMANDING OFFICER
Naval Submarine Medical Center

Approved for public release; distribution unlimited

SUMMARY PAGE

THE PROBLEM

To determine the fastest and most valid method for examining the speech intelligibility achieved by an individual or a communications system.

FINDINGS

One method recently suggested for sentence intelligibility testing is to ask the listener not to write down the words he hears, but to simply check the ordinal number of the sentence from a printed list of ten sentences. This is fast and easy to score, but a direct comparison between this method and one in which the subject actually writes down the words heard reveals that there is no correlation between the two tasks. Thus the task of sentence identification cannot be used where a valid measure of intelligibility is desired.

APPLICATION

For communications engineers involved in determining figure-of-merit for a system, and for oto-audiologists seeking to determine the speech intelligibility of which a particular person is capable.

ADMINISTRATIVE INFORMATION

This investigation was conducted as part of Bureau of Medicine and Surgery Research Work Unit M4305.08-3003D, Development of Auditory Screening Standards for Submarine/Shipboard Personnel. The present report is No. 4 on this Work Unit. It was approved for publication on 14 October '71, and designated as NavSubMedRschLab Report No. 683.

This study was supported, in part, by ONR Contract No. N00014-68-A-0197-0001 (University of Connecticut). The author worked in the Auditory Research Branch, NSMRL as a graduate assistant.

PUBLISHED BY THE NAVAL SUBMARINE MEDICAL CENTER

ABSTRACT

This study quantified the differential effects of white noise masking on the intelligibility of sentences vs the identification of a sentence from a closed set. Twenty so-called "synthetic sentences" constructed by Speaks and Jerger for intelligibility testing, but containing very little meaning, were mixed with white noise and presented to 64 Navy enlisted men who were asked, in one case, to write down as many of the words as they could understand in each sentence, and in the second case, simply to select the correct sentence from a printed list of 10 sentences. These tasks were performed with the speech material always at 60 dB SPL, but with the white noise raised in ten steps from 60 to 94 dB. The white noise tended to obscure the low-intensity high-frequency consonant discriminations generally assumed to be necessary for intelligibility, but at all comparable signal/noise ratios left relatively untouched the perception of pitch contour and other prosodic parameters characterized by high intensity and low frequency, which made sentence identification possible, even in the absence of any intelligibility whatever. A zero correlation between the two tasks revealed that sentence identification from a closed set is a task unrelated to an understanding of the words in a sentence.

DIFFERENCES BETWEEN WORD INTELLIGIBILITY AND SENTENCE IDENTIFICATION RESPONSES TO "SYNTHETIC" SENTENCES

INTRODUCTION

Speech intelligibility has been traditionally assessed in the clinic by a wide variety of speech materials, such as monosyllabic words, nonsense syllables, and rhyming consonants. The assumption is often made that intelligibility of these materials is dependent upon frequency (spectral) discriminations[1,2,5] suggest that this assumption may be too restricted and that temporal characteristics or patterns may also underlie the intelligibility of conversational speech. Sentence tests[1,6] have been used to provide the sufficient time duration necessary for the utilization of the temporal characteristics of a speech message.

Limitations of sentence tests for intelligibility measurement have been noted.[7-10] These limitations consist primarily of the use of an open message set; the unreliability of the response which is influenced by, among other things, the subject's attitude toward the test situation, his willingness to guess, and the difficulty in scoring responses. Speaks and Jerger[7] suggested a method of using sentences which reduce these limitations, consisting of having the subject simply identify the sentence from a closed set of "synthetic sentences." The synthetic sentences were constructed of random or conditional word probabilities, so that relative informational content and sentence length could be controlled.

Jerger [2, p. 230] suggested that the "filtering characteristic" of the synthetic sentence identification response is fundamentally different from that of the word intelligibility response. Intelligibility was said to require discrimination among the relatively high-frequency consonant sounds, whereas in the identification response the low-frequency vowel information can be used to produce a correct identification.

The present investigation attempts to quantify the effects of and study the basis for differential "filtering characteristics."

METHOD

1. Preparation of Test Material.

Two different synthetic sentence lists incorporating a third-order conditional probability [8, Appendix A] consisting of ten sentences each, were used in this study. Both were recorded on a Wollensak Model 1500 tape recorder, with RMS peaks at VU=0. There was a 10-sec pause on the tape between sentences. The recorded sentences were played through a clinical audiometer at a constant intensity of 60 dB sound pressure level (SPL) and for sentences 1-10 of each list, white noise at 60, 70, 80, 82, 84, 86, 88, 90, 92, and 94 dB, respectively, was electronically mixed. The sentences and noise were re-recorded together

on a single channel of a second Wollensak recorder.

2. Response Sheets.

In the first test, the intelligibility response required that the subject write down on a sheet of paper containing 10 blank lines as many words of a sentence as he could understand from the presentation of the tape-recording. While in the second test, the identification response required that the subject simply check off the number of a particular sentence out of the list of ten synthetic sentences printed on the response sheet.

3. Subjects.

Subjects were screened for normal hearing (re: Amer. Nat'l. Standards Inst., 1969) prior to testing; 64 Navy enlisted men between the ages of 17-24 years were used.

4. Test Presentation.

The subjects were tested in groups with TDH-39 earphones in MX-41 cushions on the R ear. The L ear was covered with a dummy earphone. All tapes were presented by a Wollensak Model 1500 tape recorder in a sound-absorbent room with low ambient noise.

Prior to each identification response run, each subject was made acquainted with the sentences he would later identify. The ten sentences he would hear in the identification test were randomly scrambled and presented twice in succession at an intensity level of 70 dB SPL, with no noise competition. The sentence list was then presented at the steadily increasing signal/noise (S/N) ratios.

The order of response was counter-balanced so that half of the subjects took the intelligibility test first and half the identification test. The lists were also counter-balanced so that each was used for each response an equal number of times. There was no especial time lapse between the two tests.

5. Scoring.

a) Intelligibility Test: A subject's score for any sentence was the number of correct words. Mean score for each sentence was computed and expressed as a percentage of the total possible correct words at each S/N ratio.

b) Identification Response: A subject's score was either correct or incorrect for any sentence. Per cent correct score for each sentence was the percentage of all the men who had correctly identified that sentence.

RESULTS

There was a mean correct word intelligibility score of 18.1 words (S.D.=7.07) for all 10 sentences, and a mean correct sentence identification of 7.5 sentences (S.D.=1.61).

The Pearson product-moment between the responses, utilizing all S/N ratio was .04.

Table I gives the mean per cent scores for each S/N; across all S/N's,

Table I. The mean correct word intelligibility and sentence identification responses at the signal-to-noise ratios investigated, expressed as a per cent of the total possible correct responses respectively.
NOTE: Speech always at 60 dB SPL.

Speech-to-Noise Ratio in Decibels	Per Cent Correct Word Intelligibility	Per Cent Identification Sentence Correct
0	90.5	93.5
-10	68.5	93.5
-20	49.0	91.0
-22	33.0	91.0
-24	23.0	75.0
-26	12.0	87.0
-28	12.5	78.0
-30	2.0	47.5
-32	1.0	45.0
-34	1.0	30.0
MEAN	29.2	73.1
N: 64		

the mean per cent correct word intelligibility score was 29.2, the mean per cent sentence identification score was 73.1 (t test significant at .01 level). Fig. 1 shows that the S/N at the 50% correct point was about -18 dB for word intelligibility, about -32 dB for sentence identification.

DISCUSSION

The zero correlation between the two responses to the same sentences implies that the two response measures are different procedures and are not necessarily sensitive to the same parameters of speech perception. The

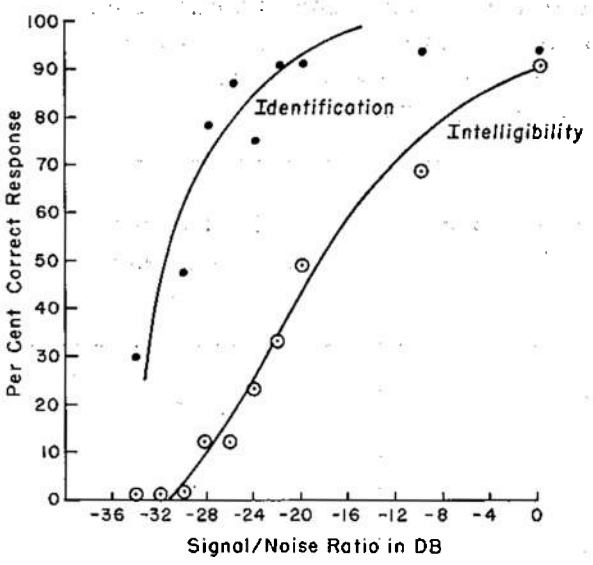


Fig. 1. Difference between the two tests at the 50% correct performance mark.

significant difference between the percentage mean correct scores for the two responses further substantiates the fundamental difference suggested by the zero correlation. Furthermore, the difference of about 14 dB between the two tests at the 50% correct performance is in general maintained at all levels of performance, as is shown by the continued separation of the curves in Fig. 1.

The basis for the differential sensitivity of the two responses to the important variables of speech perception might best be pursued along psycholinguistic lines. The separation of the curves in Fig. 1 indicates that the identification response is quite generally more resistant to white noise interference. This greater resistance may reside in the possibility that correct identification of an entire sentence, within such a closed message set, is possible with correct perception of as little as one word, or even part of a word, unique to that sentence. Further, the overall intonation pattern of each sentence is unique in a closed set. Since the subject has, then, at most, only ten intonation patterns to discriminate, it is possible, and, from the data in Fig. 1 even probable, that sentences at some S/N ratios can be identified correctly, without even one word being correctly perceived. On the other hand, word intelligibility certainly requires discrimination of the encoded phonemic elements.

A short excursion into some proposed parameters underlying the speech code may shed light here. It has been demonstrated[3,4] that encoded information about underlying consonantal phonemic structure is carried, in part, by the less intense, higher frequency second formant and its concomitant transitions. The first formant is thought to carry the prosodic features of an utterance. Because of the sharply asymmetrical upward spread of masking, high levels of white noise (with equal intensity at all frequencies) might be expected to differentially and more severely degrade the low-intensity, high-frequency second formant transitions than the more intense and lower frequency first formant. The end effect of the differential masking may have shown up in the present results. Correct prosodic feature perception, then, may be all that is needed to identify a sentence correctly in a closed message set.

It might be added that if, indeed, perception of the first formant is important to identification of a sentence in a closed set, then distorting the

temporal cues afforded by the first formant should reduce identification score; perhaps then other parameters of the speech code would take on added importance.

If we assume that consonants are important carriers of intelligibility, then the temporal parameters of the second formant transitions should be studied further. The synthetic sentences are heavily loaded with vowel information[2] and do not, therefore, depend strongly upon the perception of the second formant transitions, i.e., a person may shift almost completely to a vowel detection strategy. Since consonant detection strategies seem to be the more important in designing a measure of communicative handicap, research in and development of tests of perception strategies seem warranted.

REFERENCES

1. Davis, H., and Silverman, S. Hearing and Deafness. Holt, Rinehart, and Winston: New York, 1970.

2. Jerger, J. p. 230 in: Speech Audiometry, Rojskaer, C. (Ed.), Anelsbogtrykkeriet I Odense, Denmark, 1970. 44-65, 230, 1970.

3. Liberman, A. M., Cooper, F. S., Shankweiler, D. D. and M. Studdert-Kennedy, Perception of Speech Code, Psych. Review 74, pp 431-461, 1967.

4. Liberman, A. M. The Grammars of Speech and Language. Cognitive Psychol., 1: 1-10, 1970.

5. Pederson, O. Pp. 12-27 in: Speech Audiometry. Rojskaer, C. (Ed), Anelsbogtrykkeriet I Odense, Denmark, 1970.

6. Hudgins, C. V., Hawkins, J. E., Karlin, J. E., and Stevens, S. S. The Development of Recorded Auditory Tests for Measuring Hearing Loss for Speech. Laryngoscope, 57: 57-89, 1947.

7. Speaks, C., and Jerger, J. Method for Measurement of Speech Identification. J. Speech Hear. Res., 8: 185-194, 1965.

8. Speaks, C., Jerger, J., and Malmquist, C. Application of the Theory of Signal Detectability to Synthetic Sentence Identification. Techn. Rep. No. 7, Houston Speech and Hearing Center, 1966.

9. Giolas, T. Comparative Intelligibility Scores of Sentence Lists and Continuous Discourse. J. Audit. Res., 6: 31-38, 1966.

10. Giolas, T. Speech Audiometry. Pp. 32-56 in: Audiometry for the Retarded. Fulton R. T., and Lloyd, L. L. (Eds.) Williams and Wilkins: Baltimore, 1969.

APPENDIX A

Third-Order "Synthetic" Sentence Message Sets (From Jerger, 1970)

1. SMALL BOAT WITH A PICTURE HAS BECOME	1. OWN LOT IN YOUR AMERICAN FOOD HAS
2. BUILT THE GOVERNMENT WITH THE FORCE ALMOST	2. IS MEETING IN THE COLLEGE OF THE
3. GO CHANGE YOUR CAR COLOR IS RED	3. US WITHOUT FIRST THOUGHT WAS VERY NICE
4. FORWARD MARCH SAID THE BOY HAD A	4. AGREE WITH US IT ONLY SHALL COOK
5. MARCH AROUND WITHOUT A CARE IN YOUR	5. WITH HUMAN NATURE CAN BE MET AT
6. THAT NEIGHBOR WHO SAID BUSINESS IS BETTER	6. IN OUR DIFFERENT NAME BECAUSE HE CAN'T
7. BATTLE CRY AND BE BETTER THAN EVER	7. MARCH AROUND THE TIME WAS EVERYTHING THAT
8. DOWN BY THE TIME IS REAL ENOUGH	8. MORE OF IT IS CERTAIN TO BE
9. AGREE WITH HIM ONLY TO FIND OUT	9. WE'LL GO TO SCHOOL TODAY WAS THE
10. WOMEN VIEW MEN WITH GREEN PAPER SHOULD	10. SO ALLOW US TO THE SOUND AND

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
Submarine Medical Research Laboratory
NAVAL SUBMARINE MEDICAL CENTER

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

3. REPORT TITLE
Differences between Word Intelligibility and Sentence Identification Responses to "Synthetic" Sentences

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Interim

5. AUTHOR(S) *(First name, middle initial, last name)*
Joseph J. Smaldino and C. K. Myers

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
14 Oct 1971	6	10

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
ONR N00014-68-A-0197-0001 (Univ. Conn.)	NSMRL Report No. 683
b. PROJECT NO. BUMED M4305.08-3003D	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	

10. DISTRIBUTION STATEMENT
Approved for public release; distribution unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Naval Submarine Medical Center Box 600 Naval Submarine Base Groton, Connecticut 06340

13. ABSTRACT

This study compared directly the responses of word intelligibility and of sentence identification to the identical speech material. "Synthetic" sentences constructed by Speaks and Jerger, containing very little meaning, were mixed with white noise and presented to 64 Navy enlisted men who were asked in one case to write down as many words as they understood in each sentence, and in the second case simply to select the correct sentence from a printed list of ten sentences. The white noise tended to obscure the high-frequency consonant discriminations necessary for intelligibility, but left relatively untouched the perception of pitch contour and other prosodic parameters which made sentence identification still possible. A zero correlation between the two tasks revealed that sentence identification from a closed set of sentences is a task unrelated to an understanding of the individual words in a sentence. Thus sentence identification, though quick and useful for some purposes, cannot be used to assess intelligibility as such.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
communications system testing speech reception synthetic sentences word intelligibility sentence identification						